PATENT ABSTRACTS OF JAPAN

(11)Publication number :   **2000-227922**

(43)Date of publication of application : **15.08.2000**

(51)Int.Cl.	**G06F 17/30**

(21)Application number : **11-029501**

(22)Date of filing :   **08.02.1999**

(71)Applicant : **NIPPON TELEGR & TELEPH CORP <NTT>**
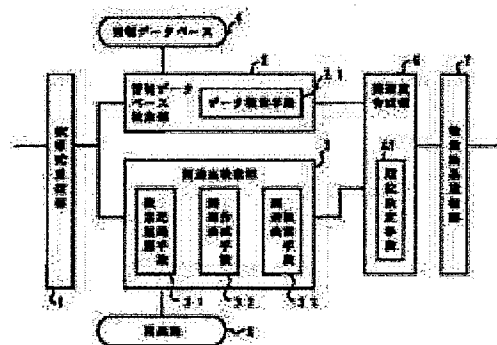
(72)Inventor : **OKUBO MASAKATSU**
SUGIZAKI MASAYUKI
MORI DAIJIRO
TANAKA KAZUO

(54) **INFORMATION RETRIEVAL DEVICE AND METHOD THEREFOR, AND RECORDING MEDIUM RECORDING PROGRAM FOR EXECUTING ITS METHOD**

(57)Abstract:

PROBLEM TO BE SOLVED: To preferentially output information which a user desires to have by calculating a point based on relation concerning retrieved data, deciding the order according to the point order and displaying information on data retrieved in accordance with the order.

SOLUTION: The data retrieval means 21 of an information data base retrieval part 2 retrieves desired data from an information data base 4 based on an inputted retrieval expression. The order deciding means 61 of an associated degree synthesis part 6 calculates a point by using a retrieval result obtained by the relation table retrieval means 33 of an relation table retrieval part 3 and decides the order in the point order. A retrieval result transmission part 7 transits the result to terminal equipment in accordance with the decided order and the terminal equipment display information on retrieved data. Thus, information which a user desires to have can preferentially be outputted and the operability of an information supply system can be improved.

CLAIMS

[Claim(s)]
[Claim 1]In an information retrieval device which searches data based on an inputted search
formula, and provides search results, A search history recording device which records data
which was retrieving time, a search formula used when searching, and the target of a detailed
presentation request at least, Relation with data which was a search term contained in said
search formula at least and the target of said detailed presentation request at a key about said
search history is collected, An association table preparing means which creates an association
table which recorded the intensive result, and a data retrieval means to search data based on
said inputted search formula, An association table search means to search relation with data
which was a search term currently recorded on said association table based on said inputted
search formula, and the target of a detailed presentation request, An information retrieval
device computing a score based on said relation about said searched data, offering the
ranking determination means which determines an order in order of a score, and displaying
information about said searched data according to said ranking.
[Claim 2]A means by which said association table preparing means classifies said input search
formula into a search term and a search condition, According to said search condition, give a
weighting factor, and it has a means to compute the sum which multiplied the number of times
of a detailed presentation request by said weighting factor for every combination of data and a
search term which were the targets of each detailed presentation request, The information
retrieval device according to claim 1 creating an association table which described said sum as
an element numerical value.
[Claim 3]The information retrieval device according to claim 2 characterized by a thing which
said ranking determination means searched an element numerical value recorded on said
association table about said decomposed search term, and were the targets of said detailed
presentation request, and for which said score is computed according to said search condition
based on said element numerical value for every data.
[Claim 4]The information retrieval device according to claim 3, wherein said ranking

determination means computes said score based on said element numerical value which carried out weighting with a using rate of each search term which was the target of said detailed presentation request in calculation of said score, and which is contained in a search formula for every data.

[Claim 5]Claims 1-4 said data retrieval means' having an inputted search formula and a means to calculate goodness of fit between each data, and said ranking determination means's amending said score using said goodness of fit, and determining said ranking with said amended score are the information retrieval devices of a description either.

[Claim 6]In an information retrieval method which searches data based on an inputted search formula, and provides search results, A search history recording process which records data which was retrieving time, a search formula used when searching, and the target of a detailed presentation request at least, Relation with data which was a search term contained in said search formula at least and the target of said detailed presentation request at a key about said search history is collected, An association table production process which creates an association table which recorded the intensive result, and a data retrieval process in which data is searched based on said inputted search formula, An association table search process in which relation with data which was a search term currently recorded on said association table based on said inputted search formula and the target of a detailed presentation request is searched, An information retrieval method computing a score based on said relation about said searched data, performing the ranking decision process which determines an order in order of a score, and displaying information about said searched data according to said ranking.

[Claim 7]The information retrieval method according to claim 6 which is provided with the following and characterized by creating an association table which described said sum as an element numerical value.

A process in which said association table production process classifies said input search formula into a search term and a search condition.

A process in which the sum which multiplied the number of times of a detailed presentation request by said weighting factor for every combination of data and a search term which gave a weighting factor according to said search condition, and were the targets of each detailed presentation request is computed.

[Claim 8]The information retrieval method according to claim 7 characterized by a thing which said ranking decision process searched an element numerical value recorded on said association table about said decomposed search term, and were the targets of said detailed presentation request, and for which said score is computed according to said search condition based on said element numerical value for every data.

[Claim 9]The information retrieval method according to claim 8, wherein said ranking decision process computes said score based on said element numerical value which carried out weighting with a using rate of each search term which was the target of said detailed

presentation request in calculation of said score, and which is contained in a search formula for every data.

[Claim 10]Claims 6-9 said data retrieval process's having an inputted search formula and a process in which goodness of fit between each data is calculated, and said ranking decision process's amending said score using said goodness of fit, and determining said ranking with said amended score are the information retrieval methods of a description either.

[Claim 11]A search history recording process which records data which was retrieving time, a search formula used when searching, and the target of a detailed presentation request at least, Relation with data which was a search term contained in said search formula at least and the target of said detailed presentation request at a key about said search history is collected, An association table production process which creates an association table which recorded the intensive result, and a data retrieval process in which data is searched based on an inputted search formula, An association table search process in which relation with data which was a search term currently recorded on said association table based on said inputted search formula and the target of a detailed presentation request is searched, A recording medium which recorded a program which computes a score based on said relation about said searched data, performs the ranking decision process which determines an order in order of a score, and performs an information retrieval method displaying information about said searched data according to said ranking.

[Claim 12]A recording medium which recorded a program which is provided with the following and characterized by creating an association table which described said sum as an element numerical value, and which performs the information retrieval method according to claim 11.
A process in which said association table production process classifies said input search formula into a search term and a search condition.
A process in which the sum which multiplied the number of times of a detailed presentation request by said weighting factor for every combination of data and a search term which gave a weighting factor according to said search condition, and were the targets of each detailed presentation request is computed.

[Claim 13]Said ranking decision process searches an element numerical value recorded on said association table about said decomposed search term, A recording medium which recorded a program which performs the information retrieval method according to claim 12 characterized by a thing which were the targets of said detailed presentation request, and for which said score is computed according to said search condition based on said element numerical value for every data.

[Claim 14]Said ranking decision process, A recording medium which recorded a program which performs the information retrieval method according to claim 13 computing said score based on said element numerical value which carried out weighting with a using rate of each search term which was the target of said detailed presentation request in calculation of said

score, and which is contained in a search formula for every data.

[Claim 15]Have said data retrieval process and an inputted search formula and a process in which goodness of fit between each data is calculated said ranking decision process, Claims 11-14 amending said score using said goodness of fit, and determining said ranking with said amended score are the recording media which recorded a program which performs an information retrieval method of a description either.

[Translation done.]

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention]When this invention creates search results using a computer network, it relates to the recording medium which recorded the program which performs an information retrieval device which carries out ranking based on the search results of the past fixed time, and is outputted, a method, and a method for the same.

[0002]

[Description of the Prior Art]Information retrieval stores the data of a document etc. in the database etc., and is art which picks out the document relevant to the search formula which the user gave from the database. With a search formula, not only like a word like "communication" but, for example like a "communication AND computer", Extraction of the document relevant to the word of both "communication" and a "computer" may be specified, or extraction of the document relevant to either word of "communication" and a "computer" may be specified like a "communication OR computer." When being beforehand given to the document as a keyword, or when the word is included ["related" here] in a document, it means both of **. [it] [it] Usually, search results output only the title of an applicable document as a title list, and if the document for which a user asks is in the title list, they will take out a document applicable by operation of detailed presentation.

[0003]In a conventional information retrieval device and method, when the suiting document was plurality, it was common to have outputted and displayed a title list in the order stored in the database. However, when many documents suited, there was a problem that it was difficult for a user to get a desired document out of search results, and a user's work became huge.

[0004]Then, for example, Document. (Donna Harman, "Ranking Algorithms",in William B.Frakes and Ricardo Baeza-Yates eds. Information Retrieval, pp.363-392,) Prentice Hall and 1992 have described how to ask for the degree of association of a search formula and each document of search results, based on the frequency in use of each search term contained in a search formula. In the database searching device of JP,04-252376,A, the freaky word storage

classified by user was provided, and search results were rearranged and displayed on order with many numbers containing a keyword. That is, solution of the above-mentioned problem has been tried by performing ranking based on the degree of association of each search term contained in a search formula, and each document in a database, and rearranging and outputting to the order.

[0005]

[Problem(s) to be Solved by the Invention]In the conventional ranking which has so far been described, if the target database does not change, the ranking of the search results to the same search formula is always the same. However, actually, even if contained in the title list, there is a document without a detailed presentation request, once there is a document which had the detailed presentation request by many users repeatedly. That is, the title list created to a certain search formula had not necessarily become a satisfying order to a user's demand.

[0006]By making this invention in view of the above problems, and creating a title list sequentially from what has many [actually] detailed presentation requests out of what was searched in the past by the same search formula as the search formula which the user inputted, It aims at the information of operativity outputted by giving priority which a user wants improving.

[0007]

[Means for Solving the Problem]In [in order to attain the above-mentioned purpose] this invention, Data which was retrieving time, a search formula used when searching, and the target of a detailed presentation request at least is recorded, Relation with data which was a search term contained in said search formula at least and the target of said detailed presentation request at a key about said search history is collected, Create an association table which recorded the intensive result, and data is searched based on said inputted search formula, Relation with data which was a search term recorded on said association table based on said inputted search formula and the target of a detailed presentation request is searched, Based on said relation, a score is computed about said searched data, an order is determined in order of the score, and it is characterized by what information about said searched data is displayed for according to the ranking.

[0008]According to this invention, a relation table which collected relation between a search formula and detailed presentation is created from a search history (a recording device, a recording process) which recorded a relation of a search formula and detailed presentation (specify and peruse data obtained as a result of search) of search which were performed during the past fixed time.

[0009]If a search formula is inputted from a terminal unit, from a created relation table, by search formula same in the past, a document with detailed presentation having been carried out [much] will be taken out in order, a title list will be created, and a user will be shown.

[0010]For this reason, since it is shown by search formula same in the past sequentially from a document which had much detailed presentation, and a document judged that many past

users have adapted themselves to that search formula most when put in another way, time until a user acquires desired information, and mitigation of a burden can be aimed at.
[0011]
[Embodiment of the Invention]Hereafter, this invention is explained in full detail based on Drawings.
[0012]Drawing 1 is a block diagram showing an example at the time of constituting this invention with a document-retrieval portion, and drawing 2 shows the flow chart of the processing.
[0013]the numerals 1 in a figure -- a search formula receive section and 2 -- an information database retrieval part and 3, as for an association table retrieval part and 4, an information database and 5 express an association table, 6 expresses a degree-of-association synchronizer, and 7 expresses the search-results transmission section.
[0014]The information database retrieval part 2 has a data retrieval means 21 to search the data for which it asks from the information database 4 based on the inputted search formula.
[0015]The association table retrieval part 3 has the search history recording device 31 which records the bibliographic data of the data which was the sent information, for example, retrieving time and the search formula used when searching, and the target of detailed presentation. The association table retrieval part 3 collects relation with the data which was a search term contained in a search formula, and the target of the detailed presentation request based on the information on the search history recording device 31, and has the association table preparing means 32 which creates the association table which summarized the intensive result. The association table retrieval part 3 has an association table search means 33 to search relation with the data which was the search term currently recorded on the association table 5 based on the inputted search formula, and the target of the detailed presentation request.
[0016]About the data searched by the data retrieval means 21, using the search results by the association table search means 33, the degree-of-association synchronizer 6 computes a score, and has the ranking determination means 61 which determines an order in order of a score.
[0017]Hereafter, it explains, doubling and referring to a flow chart.
[0018]Step (S1): Receive first the search formula sent from the terminal unit via a network in the search formula receive section 1.
[0019]The information database retrieval part 2 searches a document from the information database 4 by the search formula obtained by step (S2): (S1).
[0020]The association table retrieval part 3 searches a document using the association table 5 by the search formula obtained by step (S3): (S1).
[0021]The degree-of-association synchronizer 6 compounds the result obtained by step (S4): (S2), and the result obtained by (S3), and ranking of the search results is carried out.
[0022]The search-results transmission section 7 transmits the result obtained by step (S5):

(S4) to a terminal unit.

[0023]An order of search (S2) of a database and the search (S3) using an association table is arbitrary.

[0024]The method of creating an association table used at a step (S3) is shown below. Drawing 3 is an example of the retrieval log which shows a search term and record of detailed presentation. The 1st line means that "user U 12345" asked for detailed presentation of "the document 3" (it is a name of a document in practice) from the search results by the "Lion AND meteor stream" at "0:43 51 seconds on November 17, 1998" in drawing 3. The 2nd line or subsequent ones is the same.

[0025]Drawing 4 is a block diagram about the creation of an association table performed by calculating the relation between a search formula and a document from the retrieval log shown in drawing 3. The numerals 8 in a figure express a search formula decomposition part, 9 expresses a dignity deciding part, and 10 expresses the association table updating section.

[0026]First, the search formula decomposition part 8 decomposes each search formula into search conditions, such as a search term, AND, and OR. Next, the dignity deciding part 9 determines the dignity of each search term according to a search condition. At this example, a search formula is X_1 AND X_2 AND ... When it is AND X_N, dignity of each search term X_i (i= 1, 2, ..., N) is set to 1. A search formula is X_1 OR X_2 OR ... When it is OR X_N, dignity of each search term is set to 1/N.

[0027]It updates by applying the dignity to the **** association table showing the determined dignity in drawing 5. In drawing 5, 1 is added to the "Lion" of the line of "the document 3", and a sequence with the "meteor stream" as a processing result to the 1st line of a retrieval log, respectively.

[0028]An association table is completed by repeating the above-mentioned processing over all the periods that serve as a candidate for a total among retrieval logs. Thus, the created association table has brought a result in which many users' idea was reflected about the document suitable for each search term. Therefore, information [want / a user] can obtain the search results which the higher rank was shown by displaying search results systematically based on the association table.

[0029]As another example of the deciding method of dignity, dignity may be changed according to retrieving time. For example, an association table which thinks the more nearly latest search results as important can be created by applying the reciprocal of the difference of the date and time of creation of an association table, and the time of each search to the above-mentioned dignity. Thus, although an example of creation of an association table was shown, if the relation of a search term and a document is expressed, about the calculation method and data structure, it is arbitrary.

[0030]Hereafter, as shown in drawing 6, in an association table, the column which the line of X_i and the sequence of D_k intersect is made to describe it as $C_{i \text{ and } k}$.

[0031]The document applicable to a search formula is searched with a step (S3) using an association table, and a score is given to each document.

[0032]As a search input, when X_p OR X_q is inputted, score $W_{k[\text{ of each } D_k]}$ is made into $W_k = C_{p,k} + C_{q,k}$, and k (k= 1, 2, ...). On the other hand, as a search input, when W_p AND W_q is inputted, it is $W_k = \min(C_{p,k}, C_{q,k})$ about score $W_{k[\text{ of each } D_k]}$.

It carries out. however, $\min(a_1, a_2, ...)$ -- $a_1, a_2,$ and ... let the minimum thing in be the value. thus, each document $D_1, D_2,$ and ... the score W_1 of, $W_2,$ and ... it asks for·

[0033]As a calculation method of a score, the using rate of each search term contained in a search formula may be used. In this case, for example, as a search input, when X_p OR X_q is inputted, it is score $W_{k[\text{ of } D_k]}$ $W_k = (C_{p,k}/\text{sigma}C_{p,i}) + (C_{q,k}/\text{sigma}C_{q,i})$

It carries out. On the other hand, as a search input, when X_p AND X_q is inputted, score $W_{k[\text{ of each } D_k]}$ is set to $W_k = \min(C_{p,k}/\text{sigma}C_{p,i}) ((C_{q,k}/\text{sigma}C_{q,i}))$. thus, each document $D_1, D_2,$ and ... the score W_1 of, $W_2,$ and ... it asks for·

[0034]Although the above showed the example of the score calculation method of each document to the inputted search formula, application is arbitrary if it is a calculation method based on an association table.

[0035]In a step (S4), the result searched based on the association table and the result searched from the database are compounded.

[0036]In the search from a database, as stated to conventional technology, the goodness of fit between a search formula and each document is calculated in many cases. Here, goodness of fit which each document D_k was given is made into V_k. At this time, it asks for score $E_{k[\text{ of each document } D_k]}$ by $E_k = z \cdot (V_k/\max(V_1, V_2, ...)) + (1-z) \cdot (W_k/\max(W_1, W_2, ...))$. However, max(es) $(V_1, V_2, ...)$ are $V_1, V_2,$ and ... It is considered as the greatest inner value and referred to as $0 <= z <= 1$.

[0037]According to the upper type, at the time of z= 1, the same search results as the usual database retrieval are obtained, and when it is z= 0, the search results which used the association table are obtained. At the time of 0< z<1, the score which employed both character efficiently is generated by adding to each of the search results (usual) from a database, and the search results which used the association table, applying it. Since a score and goodness of fit serve as a value between 0 and 1 by normalizing each, the value of a score also turns into a value between 0 and 1.

[0038]At a step (S5), with the score called for by doing in this way, search results are sorted and search results are outputted. From this, the data which many users desire can always be

shown to a higher rank, and the efficiency of a user's search can be increased remarkably.

[0039]Drawing 7 shows another embodiment of the information retrieval device of this invention. The numerals 1 thru/or 7 in a figure correspond to drawing 1, and the numerals 11 are information relating parts.

[0040]In drawing 7, the information relating part 11 is formed compared with the thing of drawing 1. This information relating part 11 has the association table creation function shown in a retrieval log and drawing 4 in the form separated out of the function of the association table retrieval part 3 shown in drawing 1, and if a detailed presentation request is inputted, it will update an association table from that retrieval log. Therefore, the association table which always reflected information needs is obtained.

[0041]In the above, although the device and the method that the information which a user wants is outputted by giving priority were explained, the method can be described in the form of a program, can be stored in a recording medium, and can be kept refreshable. Therefore, this invention also contains the recording medium concerned in technical scope.

[0042]

[Effect of the Invention]Since according to this invention ranking attachment of the search results is carried out automatically and they are outputted based on the association table obtained by analyzing the retrieval log of the prescribed period which many users used, as explained above, The information which a user wants has priority, and is outputted and it is effective in the operativity of an information service system improving extremely.

[Translation done.]

TECHNICAL FIELD

[Field of the Invention]When this invention creates search results using a computer network, it relates to the recording medium which recorded the program which performs an information retrieval device which carries out ranking based on the search results of the past fixed time, and is outputted, a method, and a method for the same.

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any

damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

PRIOR ART

[Description of the Prior Art]Information retrieval stores the data of a document etc. in the database etc., and is art which picks out the document relevant to the search formula which the user gave from the database. With a search formula, not only like a word like "communication" but, for example like a "communication AND computer", Extraction of the document relevant to the word of both "communication" and a "computer" may be specified, or extraction of the document relevant to either word of "communication" and a "computer" may be specified like a "communication OR computer." When being beforehand given to the document as a keyword, or when the word is included ["related" here] in a document, it means both of **. [it] [it] Usually, search results output only the title of an applicable document as a title list, and if the document for which a user asks is in the title list, they will take out a document applicable by operation of detailed presentation.

[0003]In a conventional information retrieval device and method, when the suiting document was plurality, it was common to have outputted and displayed a title list in the order stored in the database. However, when many documents suited, there was a problem that it was difficult for a user to get a desired document out of search results, and a user's work became huge.

[0004]Then, for example, Document. (Donna Harman, "Ranking Algorithms",in William B.Frakes and Ricardo Baeza-Yates eds. Information Retrieval, pp.363-392,) Prentice Hall and 1992 have described how to ask for the degree of association of a search formula and each document of search results, based on the frequency in use of each search term contained in a search formula. In the database searching device of JP,04-252376,A, the freaky word storage classified by user was provided, and search results were rearranged and displayed on order with many numbers containing a keyword. That is, solution of the above-mentioned problem has been tried by performing ranking based on the degree of association of each search term contained in a search formula, and each document in a database, and rearranging and outputting to the order.

[Translation done.]

EFFECT OF THE INVENTION

[Effect of the Invention]Since according to this invention ranking attachment of the search results is carried out automatically and they are outputted based on the association table obtained by analyzing the retrieval log of the prescribed period which many users used, as explained above, The information which a user wants has priority, and is outputted and it is effective in the operativity of an information service system improving extremely.

[Translation done.]

TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention]In the conventional ranking which has so far been
described, if the target database does not change, the ranking of the search results to the
same search formula is always the same. However, actually, even if contained in the title list,
there is a document without a detailed presentation request, once there is a document which
had the detailed presentation request by many users repeatedly. That is, the title list created to
a certain search formula had not necessarily become a satisfying order to a user's demand.
[0006]By making this invention in view of the above problems, and creating a title list
sequentially from what has many [actually] detailed presentation requests out of what was
searched in the past by the same search formula as the search formula which the user
inputted, It aims at the information of operativity outputted by giving priority which a user wants
improving.

[Translation done.]

* NOTICES *

MEANS

[Means for Solving the Problem]In [in order to attain the above-mentioned purpose] this
invention, Data which was retrieving time, a search formula used when searching, and the
target of a detailed presentation request at least is recorded, Relation with data which was a
search term contained in said search formula at least and the target of said detailed
presentation request at a key about said search history is collected, Create an association
table which recorded the intensive result, and data is searched based on said inputted search
formula, Relation with data which was a search term recorded on said association table based
on said inputted search formula and the target of a detailed presentation request is searched,
Based on said relation, a score is computed about said searched data, an order is determined
in order of the score, and it is characterized by what information about said searched data is
displayed for according to the ranking.
[0008]According to this invention, a relation table which collected relation between a search
formula and detailed presentation is created from a search history (a recording device, a
recording process) which recorded a relation of a search formula and detailed presentation
(specify and peruse data obtained as a result of search) of search which were performed
during the past fixed time.
[0009]If a search formula is inputted from a terminal unit, from a created relation table, by
search formula same in the past, a document with detailed presentation having been carried
out [much] will be taken out in order, a title list will be created, and a user will be shown.
[0010]For this reason, since it is shown by search formula same in the past sequentially from a
document which had much detailed presentation, and a document judged that many past
users have adapted themselves to that search formula most when put in another way, time
until a user acquires desired information, and mitigation of a burden can be aimed at.
[0011]
[Embodiment of the Invention]Hereafter, this invention is explained in full detail based on
Drawings.
[0012]Drawing 1 is a block diagram showing an example at the time of constituting this

invention with a document-retrieval portion, and drawing 2 shows the flow chart of the processing.

[0013]the numerals 1 in a figure -- a search formula receive section and 2 -- an information database retrieval part and 3, as for an association table retrieval part and 4, an information database and 5 express an association table, 6 expresses a degree-of-association synchronizer, and 7 expresses the search-results transmission section.

[0014]The information database retrieval part 2 has a data retrieval means 21 to search the data for which it asks from the information database 4 based on the inputted search formula.

[0015]The association table retrieval part 3 has the search history recording device 31 which records the bibliographic data of the data which was the sent information, for example, retrieving time and the search formula used when searching, and the target of detailed presentation. The association table retrieval part 3 collects relation with the data which was a search term contained in a search formula, and the target of the detailed presentation request based on the information on the search history recording device 31, and has the association table preparing means 32 which creates the association table which summarized the intensive result. The association table retrieval part 3 has an association table search means 33 to search relation with the data which was the search term currently recorded on the association table 5 based on the inputted search formula, and the target of the detailed presentation request.

[0016]About the data searched by the data retrieval means 21, using the search results by the association table search means 33, the degree-of-association synchronizer 6 computes a score, and has the ranking determination means 61 which determines an order in order of a score.

[0017]Hereafter, it explains, doubling and referring to a flow chart.

[0018]Step (S1): Receive first the search formula sent from the terminal unit via a network in the search formula receive section 1.

[0019]The information database retrieval part 2 searches a document from the information database 4 by the search formula obtained by step (S2): (S1).

[0020]The association table retrieval part 3 searches a document using the association table 5 by the search formula obtained by step (S3): (S1).

[0021]The degree-of-association synchronizer 6 compounds the result obtained by step (S4): (S2), and the result obtained by (S3), and ranking of the search results is carried out.

[0022]The search-results transmission section 7 transmits the result obtained by step (S5): (S4) to a terminal unit.

[0023]An order of search (S2) of a database and the search (S3) using an association table is arbitrary.

[0024]The method of creating an association table used at a step (S3) is shown below. Drawing 3 is an example of the retrieval log which shows a search term and record of detailed presentation. The 1st line means that "user U 12345" asked for detailed presentation of "the

document 3" (it is a name of a document in practice) from the search results by the "Lion AND meteor stream" at "0:43 51 seconds on November 17, 1998" in drawing 3. The 2nd line or subsequent ones is the same.

[0025]Drawing 4 is a block diagram about the creation of an association table performed by calculating the relation between a search formula and a document from the retrieval log shown in drawing 3. The numerals 8 in a figure express a search formula decomposition part, 9 expresses a dignity deciding part, and 10 expresses the association table updating section.

[0026]First, the search formula decomposition part 8 decomposes each search formula into search conditions, such as a search term, AND, and OR. Next, the dignity deciding part 9 determines the dignity of each search term according to a search condition. At this example, a search formula is X_1 AND X_2 AND ... When it is AND X_N, dignity of each search term X_i (i= 1, 2, ..., N) is set to 1. A search formula is X_1 OR X_2 OR ... When it is OR X_N, dignity of each search term is set to 1/N.

[0027]It updates by applying the dignity to the **** association table showing the determined dignity in drawing 5. In drawing 5, 1 is added to the "Lion" of the line of "the document 3", and a sequence with the "meteor stream" as a processing result to the 1st line of a retrieval log, respectively.

[0028]An association table is completed by repeating the above-mentioned processing over all the periods that serve as a candidate for a total among retrieval logs. Thus, the created association table has brought a result in which many users' idea was reflected about the document suitable for each search term. Therefore, information [want / a user] can obtain the search results which the higher rank was shown by displaying search results systematically based on the association table.

[0029]As another example of the deciding method of dignity, dignity may be changed according to retrieving time. For example, an association table which thinks the more nearly latest search results as important can be created by applying the reciprocal of the difference of the date and time of creation of an association table, and the time of each search to the above-mentioned dignity. Thus, although an example of creation of an association table was shown, if the relation of a search term and a document is expressed, about the calculation method and data structure, it is arbitrary.

[0030]Hereafter, as shown in drawing 6, in an association table, the column which the line of X_i and the sequence of D_k intersect is made to describe it as $C_{i \text{ and } k}$.

[0031]The document applicable to a search formula is searched with a step (S3) using an association table, and a score is given to each document.

[0032]As a search input, when X_p OR X_q is inputted, score $W_{k[\text{ of each } D_k]}$ is made into $W_k = C_{p, k} + C_{q, \text{ and } k}$ (k= 1, 2, ...). On the other hand, as a search input, when W_p AND W_q is inputted, it

is $W_k = \min(C_{p,k}, C_{q,k})$ about score $W_{k[\text{ of each } D_k]}$.

It carries out. however, $\min(a_1, a_2, \ldots) - a_1, a_2$, and ... let the minimum thing in be the value. thus, each document D_1, D_2, and ... the score $W_{1 \text{ of }}, W_2$, and ... it asks for.

[0033]As a calculation method of a score, the using rate of each search term contained in a search formula may be used. In this case, for example, as a search input, when X_p OR X_q is inputted, it is score $W_{k[\text{ of } D_k]}$ $W_k = (C_{p,k}/\text{sigma}C_{p,i}) + (C_{q,k}/\text{sigma}C_{q,i})$

It carries out. On the other hand, as a search input, when X_p AND X_q is inputted, score $W_{k[\text{ of each } D_k]}$ is set to $W_k = \min(C_{p,k}/\text{sigma}C_{p,i})((C_{q,k}/\text{sigma}C_{q,i}))$. thus, each document D_1, D_2, and ... the score $W_{1 \text{ of }}, W_2$, and ... it asks for.

[0034]Although the above showed the example of the score calculation method of each document to the inputted search formula, application is arbitrary if it is a calculation method based on an association table.

[0035]In a step (S4), the result searched based on the association table and the result searched from the database are compounded.

[0036]In the search from a database, as stated to conventional technology, the goodness of fit between a search formula and each document is calculated in many cases. Here, goodness of fit which each document D_k was given is made into V_k. At this time, it asks for score $E_{k[\text{ of each document } D_k]}$ by $E_k = z - (V_k/\max(V_1, V_2, \ldots)) + (1-z) - (W_k/\max(W_1, W_2, \ldots))$. However, max(es) $(V_1, V_2, \ldots)$ are V_1, V_2, and ... It is considered as the greatest inner value and referred to as $0 <= z <= 1$.

[0037]According to the upper type, at the time of z= 1, the same search results as the usual database retrieval are obtained, and when it is z= 0, the search results which used the association table are obtained. At the time of 0< z<1, the score which employed both character efficiently is generated by adding to each of the search results (usual) from a database, and the search results which used the association table, applying it. Since a score and goodness of fit serve as a value between 0 and 1 by normalizing each, the value of a score also turns into a value between 0 and 1.

[0038]At a step (S5), with the score called for by doing in this way, search results are sorted and search results are outputted. From this, the data which many users desire can always be shown to a higher rank, and the efficiency of a user's search can be increased remarkably.

[0039]Drawing 7 shows another embodiment of the information retrieval device of this invention. The numerals 1 thru/or 7 in a figure correspond to drawing 1, and the numerals 11 are information relating parts.

[0040]In drawing 7, the information relating part 11 is formed compared with the thing of

drawing 1. This information relating part 11 has the association table creation function shown in a retrieval log and drawing 4 in the form separated out of the function of the association table retrieval part 3 shown in drawing 1, and if a detailed presentation request is inputted, it will update an association table from that retrieval log. Therefore, the association table which always reflected information needs is obtained.

[0041]In the above, although the device and the method that the information which a user wants is outputted by giving priority were explained, the method can be described in the form of a program, can be stored in a recording medium, and can be kept refreshable. Therefore, this invention also contains the recording medium concerned in technical scope.

[Translation done.]

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1]It is a block diagram showing an example at the time of constituting this invention with a document-retrieval portion.
[Drawing 2]It is a flow chart explaining processing of this invention.
[Drawing 3]An example of the retrieval log which shows the relation between a search term and detailed presentation is shown.
[Drawing 4]It is a block diagram which creates an association table from the retrieval log of drawing 3.
[Drawing 5]It is a figure explaining the correspondence relation between a search term and a document.
[Drawing 6]It is a figure explaining the sign assigned to each line and each sequence of the association table.
[Drawing 7]It is a block diagram showing the composition of another embodiment of this invention.
[Description of Notations]
1 Search formula receive section
2 Information database retrieval part
3 Association table retrieval part
4 Information database
5 Association table
6 Degree-of-association synchronizer
7 Search-results transmission section
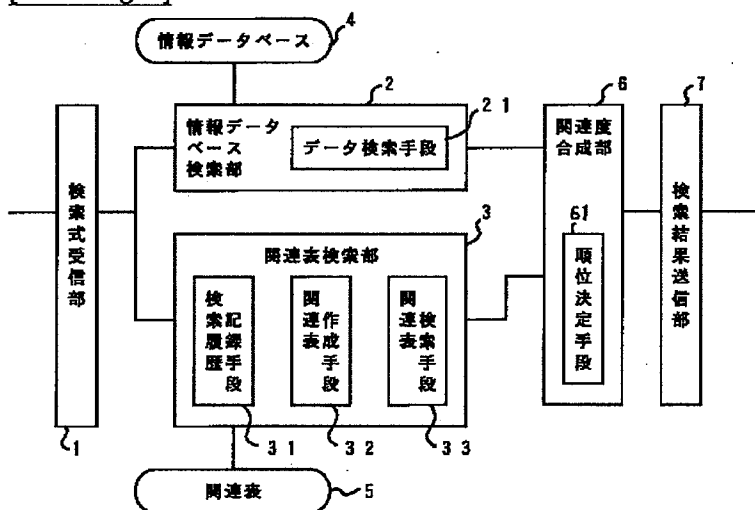11 Information relating part

[Translation done.]

* NOTICES *

JPO and INPIT are not responsible for any
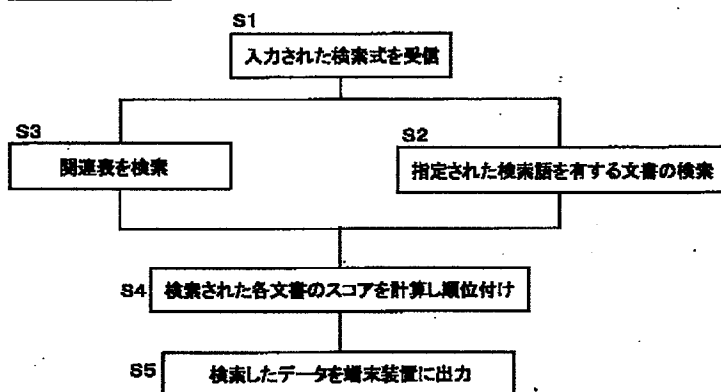
damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

DRAWINGS

[Drawing 1]

[Drawing 2]

[Drawing 3]

```
1998-11-17,00:43:51,U12345,しし座*流星群,文書3
1998-11-17,00:43:51,U12798,netaddress*登録,文書128
1998-11-17,00:43:56,U10195,しし座+獅子座,文書3
1998-11-17,00:44:00,U00285,ニセコ,文書42
1998-11-17,00:44:03,U62745,宿泊施設+ホテル,文書989
1998-11-17,00:44:05,U12900,論文*マーケティング,文書1098
1998-11-17,00:44:20,U12345,獅子座流星群*撮影,文書4098
...
```
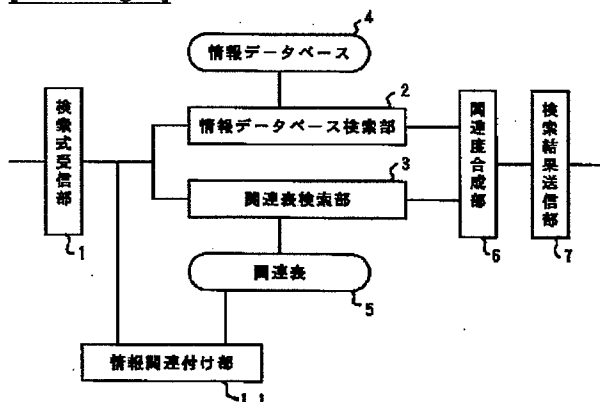
[Drawing 4]

検索式分解部	～8
重み決定部	～9
関連表更新部	～10

[Drawing 5]

文書 / 検索語	1	2	3	4	...	合計
:					...	
しし座			1		...	1
:					...	
流星群			1		...	1
:						
合計			2		...	

[Drawing 6]

文書 / 検索語	D_1	D_2	...	D_k	...	合計
X_1	$c_{1,1}$	$c_{1,2}$	...	$c_{1,k}$	...	$\sum_j c_{1,j}$
X_2	$c_{2,1}$	$c_{2,2}$	...	$c_{2,k}$	...	$\sum_j c_{2,j}$
:			...		...	
X_i	$c_{i,1}$	$c_{i,2}$	...	$c_{i,k}$	...	$\sum_j c_{i,j}$
:						
合計	$\sum_j c_{j,1}$	$\sum_j c_{j,2}$	...	$\sum_j c_{j,k}$	...	

[Drawing 7]

[Translation done.]

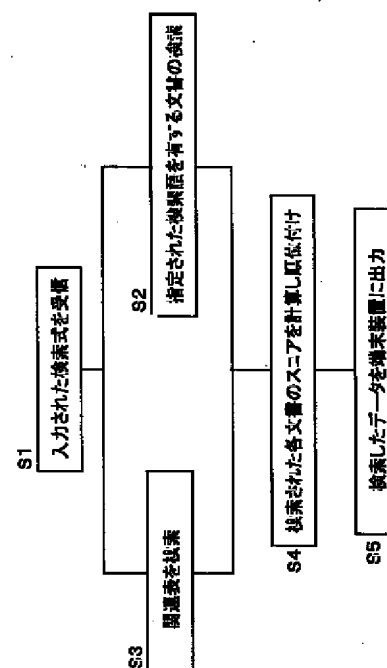| (51)Int.Cl.⁷	識別記号	F I			テーマコード（参考）
G06F　17/30		C06F	15/403	340B	5B075
			15/40	310F	
			15/403	370Z	

(54)【発明の名称】　情報検索装置、方法及びその方法を実行するプログラムを記録した記録媒体

(57)【要約】

【課題】　本発明は、過去の検索において要求の多かった情報を優先して出力できるようにすることを目的としている。

【解決手段】　検索履歴を手がかりに、過去の検索に当っての検索語と詳細提示要求のあったデータとの関連を集約して関連表を作成しておき、検索されたデータの中から関連表を用いて得られた得点の高いものをより先に出力するようにする。

【特許請求の範囲】

【請求項1】 入力された検索式に基づいてデータを検索して、検索結果を提供する情報検索装置において、
少なくとも検索時刻と検索に際して用いられた検索式と詳細提示要求の対象となったデータとを記録する検索履歴記録手段と、
前記検索履歴を手がかりに、少なくとも前記検索式に含まれる検索語と前記詳細提示要求の対象となったデータとの関連を集約して、その集約結果を記録した関連表を作成する関連表作成手段と、
前記入力された検索式をもとにデータを検索するデータ検索手段と、
前記入力された検索式をもとに前記関連表に記録されている検索語と詳細提示要求の対象となったデータとの関連を検索する関連表検索手段と、
前記検索されたデータについて前記関連に基づいて得点を算出し、その得点順に順序を決定する順位決定手段とをそなえ、
前記順位に従って前記検索されたデータに関する情報を表示することを特徴とする情報検索装置。

【請求項2】 前記関連表作成手段は、
前記入力検索式を検索語と検索条件に分類する手段と、
前記検索条件に応じて重み係数を与え、各詳細提示要求の対象となったデータと検索語との組合わせごとに詳細提示要求回数に前記重み係数を乗じた和を算出する手段とを有し、
前記和を要素数値として記述した関連表を作成することを特徴とする請求項1記載の情報検索装置。

【請求項3】 前記順位決定手段は、前記分解された検索語について前記関連表に記録された要素数値を検索し、前記詳細提示要求の対象となったデータ毎に前記要素数値を基に前記検索条件に応じて前記得点を算出することを特徴とする請求項2記載の情報検索装置。

【請求項4】 前記順位決定手段は、前記得点の算出において前記詳細提示要求の対象となったデータ毎に検索式に含まれる各検索語の使用割合で重みづけした前記要素数値を基に前記得点を算出することを特徴とする請求項3記載の情報検索装置。

【請求項5】 前記データ検索手段は、入力された検索式と各データ間における適合度を計算する手段を有し、
前記順位決定手段は、前記得点を前記適合度を用いて補正し、前記補正された得点をもって前記順位を決定することを特徴とする請求項1乃至4のいずれか記載の情報検索装置。

【請求項6】 入力された検索式に基づいてデータを検索して、検索結果を提供する情報検索方法において、
少なくとも検索時刻と検索に際して用いられた検索式と詳細提示要求の対象となったデータとを記録する検索履歴記録過程と、
前記検索履歴を手がかりに、少なくとも前記検索式に含

まれる検索語と前記詳細提示要求の対象となったデータとの関連を集約して、その集約結果を記録した関連表を作成する関連表作成過程と、
前記入力された検索式をもとにデータを検索するデータ検索過程と、
前記入力された検索式をもとに前記関連表に記録されている検索語と詳細提示要求の対象となったデータとの関連を検索する関連表検索過程と、
前記検索されたデータについて前記関連に基づいて得点を算出し、その得点順に順序を決定する順位決定過程とを実行し、
前記順位に従って前記検索されたデータに関する情報を表示することを特徴とする情報検索方法。

【請求項7】 前記関連表作成過程は、
前記入力検索式を検索語と検索条件に分類する過程と、
前記検索条件に応じて重み係数を与え、各詳細提示要求の対象となったデータと検索語との組合わせごとに詳細提示要求回数に前記重み係数を乗じた和を算出する過程とを有し、
前記和を要素数値として記述した関連表を作成することを特徴とする請求項6記載の情報検索方法。

【請求項8】 前記順位決定過程は、前記分解された検索語について前記関連表に記録された要素数値を検索し、前記詳細提示要求の対象となったデータ毎に前記要素数値を基に前記検索条件に応じて前記得点を算出することを特徴とする請求項7記載の情報検索方法。

【請求項9】 前記順位決定過程は、前記得点の算出において前記詳細提示要求の対象となったデータ毎に検索式に含まれる各検索語の使用割合で重みづけした前記要素数値を基に前記得点を算出することを特徴とする請求項8記載の情報検索方法。

【請求項10】 前記データ検索過程は、入力された検索式と各データ間における適合度を計算する過程を有し、
前記順位決定過程は、前記得点を前記適合度を用いて補正し、前記補正された得点をもって前記順位を決定することを特徴とする請求項6乃至9のいずれか記載の情報検索方法。

【請求項11】 少なくとも検索時刻と検索に際して用いられた検索式と詳細提示要求の対象となったデータとを記録する検索履歴記録過程と、
前記検索履歴を手がかりに、少なくとも前記検索式に含まれる検索語と前記詳細提示要求の対象となったデータとの関連を集約して、その集約結果を記録した関連表を作成する関連表作成過程と、
入力された検索式をもとにデータを検索するデータ検索過程と、
前記入力された検索式をもとに前記関連表に記録されている検索語と詳細提示要求の対象となったデータとの関連を検索する関連表検索過程と、

前記検索されたデータについて前記関連に基づいて得点を算出し、その得点順に順序を決定する順位決定過程とを実行し、

前記順位に従って前記検索されたデータに関する情報を表示することを特徴とする情報検索方法を実行するプログラムを記録した記録媒体。

【請求項１２】　前記関連表作成過程は、

前記入力検索式を検索語と検索条件に分類する過程と、前記検索条件に応じて重み係数を与え、各詳細提示要求の対象となったデータと検索語との組合わせごとに詳細提示要求回数に前記重み係数を乗じた和を算出する過程とを有し、

前記和を要素数値として記述した関連表を作成することを特徴とする請求項１１記載の情報検索方法を実行するプログラムを記録した記録媒体。

【請求項１３】　前記順位決定過程は、前記分解された検索語について前記関連表に記録された要素数値を検索し、前記詳細提示要求の対象となったデータ毎に前記要素数値を基に前記検索条件に応じて前記得点を算出することを特徴とする請求項１２記載の情報検索方法を実行するプログラムを記録した記録媒体。

【請求項１４】　前記順位決定過程は、前記得点の算出において前記詳細提示要求の対象となったデータ毎に検索式に含まれる各検索語の使用割合で重みづけした前記要素数値を基に前記得点を算出することを特徴とする請求項１３記載の情報検索方法を実行するプログラムを記録した記録媒体。

【請求項１５】　前記データ検索過程は、入力された検索式と各データ間における適合度を計算する過程を有し、

前記順位決定過程は、前記得点を前記適合度を用いて補正し、前記補正された得点をもって前記順位を決定することを特徴とする請求項１１乃至１４のいずれか記載の情報検索方法を実行するプログラムを記録した記録媒体。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、コンピュータネットワークを用いて検索結果を作成する際に、過去の一定期間の検索結果に基づいて順位付けをして出力する情報検索装置、方法及びその方法を実行するプログラムを記録した記録媒体に関する。

【０００２】

【従来の技術】情報検索とは、文書などのデータをデータベースなどに蓄積しておき、利用者が与えた検索式に関連する文書をそのデータベースから取り出す技術である。検索式とは、例えば「通信」のような単語だけではなく、「通信ＡＮＤ計算機」のように、「通信」と「計算機」との両方の単語に関連する文書の取り出しを指定したり、「通信ＯＲ計算機」のように、「通信」と「計算機」とのいずれかの単語に関連する文書の取り出しを指定することもある。ここで「関連する」とは、キーワードとして予め文書に付与されている場合、あるいは文書中にその語が含まれる場合、の両者を意味している。通常検索結果は、該当する文書の標題のみを標題リストとして出力し、利用者の所望する文書がその標題リストの中にあれば、詳細提示という操作によって該当する文書を取り出す。

【０００３】従来の情報検索装置および方法において、適合する文書が複数の場合、データベースに格納されている順序で標題リストを出力・表示するのが一般的であった。しかし、多数の文書が適合した場合には、検索結果の中から利用者が所望の文書を得ることが困難で、利用者の作業が膨大になるという問題があった。

【０００４】そこで、例えば、文献（Donna Harman, "Ranking Algorithms",in William B.Frakes and Ricardo Baeza-Yates eds. Information Retrieval, pp.363-392, Prentice Hall, 1992）では、検索式に含まれる各検索語の使用頻度に基づいて、検索式と検索結果の各文書との関連度を求める方法について述べている。また、特開平04-252376のデータベース検索装置では、ユーザ別フリーキーワード格納部を設け、キーワードを含有する数の多い順に検索結果を並び替えて表示していた。すなわち、検索式に含まれる各検索語とデータベース内の各文書との関連度に基づいて順位付けを行い、その順に並び替えて出力することによって、上記問題の解決を試みてきた。

【０００５】

【発明が解決しようとする課題】これまで述べてきたような従来の順位付けでは、対象となるデータベースが変わらなければ、同一の検索式に対する検索結果の順位は常に同じである。しかし実際には、標題リストに含まれていても、一度も詳細提示要求のない文書もあれば、多数の利用者によって何度も詳細提示要求のあった文書もある。すなわち、ある検索式に対して作成された標題リストは、利用者の要求に対して、必ずしも満足のいく順序となっているわけではなかった。

【０００６】本発明は上記のような問題点に鑑みてなされたものであり、利用者が入力した検索式と同じ検索式によって過去に検索されたものの中から、実際に詳細提示要求の多かったものから順に標題リストを作成することによって、利用者の欲する情報が優先して出力される、操作性の向上することを目的としている。

【０００７】

【課題を解決するための手段】上記目的を達成するため、本発明においては、少なくとも検索時刻と検索に際して用いられた検索式と詳細提示要求の対象となったデータとを記録するようにし、前記検索履歴を手がかりに、少なくとも前記検索式に含まれる検索語と前記詳細提示要求の対象となったデータとの関連を集約して、そ

の集約結果を記録した関連表を作成するようにし、前記入力された検索式をもとにデータを検索するようにし、前記入力された検索式をもとに前記関連表に記録された検索語と詳細提示要求の対象となったデータとの関連を検索するようにし、前記検索されたデータについて前記関連に基づいて得点を算出し、その得点順に順序を決定して、その順位に従って前記検索されたデータに関する情報を表示するようにする、ことを特徴としている。

【０００８】本発明によれば、過去の一定期間中に行われた検索の、検索式と詳細提示（検索の結果得られたデータを特定して閲覧すること）との関係を記録した検索履歴（記録手段、記録過程）から、検索式と詳細提示との関係を集約した関係表を作成しておく。

【０００９】端末装置から検索式が入力されると、作成された関係表から、過去に同一の検索式によって詳細提示されたことの多い文書を順に取り出して標題リストを作成し、利用者に提示する。

【００１０】このため、過去に同一の検索式によって詳細提示が多かった文書、換言すると、過去の多くの利用者が、その検索式に最も適合していると判断された文書から順に提示されるため、利用者が所望の情報を得るまでの時間、負担の軽減が図れる。

【００１１】
【発明の実施の形態】以下、本発明を図面に基づいて詳述する。

【００１２】図１は本発明を文書検索部分とともに構成した場合の一例を示すブロック図で、図２はその処理のフローチャートを示している。

【００１３】図中の符号１は検索式受信部、２は情報データベース検索部、３は関連表検索部、４は情報データベース、５は関連表、６は関連度合成部、７は検索結果送信部を表わしている。

【００１４】なお、情報データベース検索部２は、入力された検索式をもとに情報データベース４から所望されるデータを検索するデータ検索手段２１を有する。

【００１５】関連表検索部３は、送られてきた情報、例えば検索時刻と検索に際して用いられる検索式と詳細提示の対象となったデータの書誌事項を記録する検索履歴記録手段３１を持っている。また、関連表検索部３は、検索履歴記録手段３１の情報にもとづいて、検索式に含まれる検索語と詳細提示要求の対象となったデータとの関連を集約して、その集約結果をまとめた関連表を作成する関連表作成手段３２を持っている。更に関連表検索部３は、入力された検索式をもとに関連表５に記録されている検索語と詳細提示要求の対象となったデータとの関連を検索する関連表検索手段３３を有する。

【００１６】関連度合成部６は、データ検索手段２１によって検索されたデータについて、関連表検索手段３３による検索結果を利用して、得点を算出し、その得点順に順序を決定する順位決定手段６１を有する。

【００１７】以下、フローチャートを合わせ参照しつつ説明する。

【００１８】ステップ（Ｓ１）：まず、ネットワークを介して端末装置から送られてきた検索式を検索式受信部１で受信する。

【００１９】ステップ（Ｓ２）：（Ｓ１）で得られた検索式によって情報データベース検索部２が情報データベース４から文書を検索する。

【００２０】ステップ（Ｓ３）：（Ｓ１）で得られた検索式によって関連表検索部３が関連表５を用いて文書を検索する。

【００２１】ステップ（Ｓ４）：（Ｓ２）で得られた結果と、（Ｓ３）で得られた結果とを関連度合成部６が合成して検索結果を順位付けする。

【００２２】ステップ（Ｓ５）：（Ｓ４）で得られた結果を検索結果送信部７が端末装置に送信する。

【００２３】なお、データベースの検索（Ｓ２）と関連表を用いた検索（Ｓ３）との順序は任意である。

【００２４】ステップ（Ｓ３）で用いる関連表の作成法を以下に示す。図３は、検索語と詳細提示の記録とを示す検索ログの一例である。図３では、１行目は、「1998年11月17日の０時43分51秒」に、「ユーザＵ 12345」が、「しし座ＡＮＤ流星群」による検索結果から「文書３」（実際は文書の名前である）の詳細提示を求めたことを表している。２行目以降も同様である。

【００２５】図４は、図３に示した検索ログから、検索式と文書との関係を計算して行う関連表の作成についてのブロック図である。図中の符号８は検索式分解部、９は重み決定部、１０は関連表更新部を表わしている。

【００２６】まず、検索式分解部８が各検索式を検索語とＡＮＤやＯＲなどの検索条件に分解する。次に重み決定部９は検索条件に応じて各検索語の重みを決定する。本実施例では、検索式が
X_1 ＡＮＤ X_2 ＡＮＤ ... ＡＮＤ X_N
のときには、各検索語X_i（$i=1,2,...,N$）の重みを１とする。検索式が、
X_1 ＯＲ X_2 ＯＲ ... ＯＲ X_N
のときには、各検索語の重みを１／Ｎとする。

【００２７】決定された重みを、図５に示す如き関連表にその重みを加えることによって更新する。図５では、検索ログの１行めに対する処理結果として、「文書３」の行の「しし座」と「流星群」との列にそれぞれ１を加えている。

【００２８】検索ログのうち、集計対象となる期間のすべてにわたって上記処理を繰り返すことによって関連表を完成させる。このようにして作成された関連表は、各検索語に適した文書について、多くの利用者の考えが反映された結果となっている。したがって、その関連表に基づいて検索結果を順序付けて表示することにより、利用者が欲しい情報が上位に提示された検索結果を得るこ

とができる。

【００２９】重みの決定方法の別の例としては、検索時刻に応じて重みを変化させてもよい。例えば、関連表の作成日時と各検索の日時との差分の逆数を上記重みに対してかけることにより、より直近の検索結果を重視するような関連表を作成することができる。このように、関連表の作成の一例を示したが、検索語と文書との関連を表すものであれば、その計算方法やデータ構造については任意である。

【００３０】以下、図６に示すように、関連表において、X_i の行と D_k の列の交差する欄を、$C_{i,k}$ と記すことにする。

【００３１】ステップ（Ｓ３）では、関連表を用いて、検索式に該当する文書を検索し、各文書に得点を与える。

【００３２】検索入力として、X_p OR X_q が入力された場合には、各 D_k の得点 W_k を
$$W_k = C_{p,k} + C_{q,k}$$
とする（$k=1,2,...$）。一方、検索入力として、W_p AND W_q が入力された場合には、各 D_k の得点 W_k を、
$$W_k = \min(C_{p,k}, C_{q,k})$$
とする。ただし、$\min(a_1, a_2, ...)$ は、$a_1, a_2, ...$ の中で最小のものをその値とする。このようにして、各文書 $D_1, D_2, ...$ の得点 $W_1, W_2, ...$ を求める。

【００３３】得点の計算方法として、検索式に含まれる各検索語の使用割合を用いてもよい。この場合、例えば、検索入力として、X_p OR X_q が入力されたとき、D_k の得点 W_k を、
$$W_k = (C_{p,k}/\Sigma C_{p,i}) + (C_{q,k}/\Sigma C_{q,i})$$
とする。一方、検索入力として、X_p AND X_q が入力された場合には、各 D_k の得点 W_k を、
$$W_k = \min((C_{p,k}/\Sigma C_{p,i}), (C_{q,k}/\Sigma C_{q,i}))$$
とする。このようにして、各文書 $D_1, D_2, ...$ の得点 $W_1, W_2, ...$ を求める。

【００３４】以上は、入力された検索式に対する各文書の得点計算方法の例を示したが、関連表に基づいた計算方法であれば適用は任意である。

【００３５】ステップ（Ｓ４）では、関連表に基づいて検索された結果と、データベースから検索された結果とを合成する。

【００３６】データベースからの検索では、従来技術に述べたように、検索式と各文書との間の適合度を計算することが多い。ここでは、各文書 D_k に与えられた適合度を V_k とする。このとき、各文書 D_k のスコア E_k を、
$$E_k = z \cdot (V_k/\max(V_1, V_2, ...)) + (1-z) \cdot (W_k/\max(W_1, W_2, ...))$$
によって求める。ただし、$\max(V_1, V_2, ...)$ は、$V_1, V_2, ...$ のうちの最大の値とし、$0 \leq z \leq 1$ とする。

【００３７】上式によれば、$z=1$ のときには、通常のデータベース検索と同様の検索結果が得られ、$z=0$ のときには関連表を用いた検索結果が得られる。$0<z<1$ のときには、データベースからの（通常の）検索結果と、関連表を用いた検索結果のそれぞれに重みをかけて加えることによって、両者の性質を生かしたスコアを生成する。また、それぞれを正規化することにより、得点、および適合度は０から１の間の値となるため、スコアの値も０から１の間の値となる。

【００３８】ステップ（Ｓ５）ではこのようにして求められたスコアによって検索結果をソートし、検索結果を出力する。このことから、多くの利用者が望んでいるデータを常に上位に提示することができ、利用者の検索を著しく効率化できる。

【００３９】図７は本発明の情報検索装置の別の実施形態を示している。図中の符号１ないし７は図１に対応し、符号１１は情報関連付け部である。

【００４０】図７においては図１のものにくらべて情報関連付け部１１が設けられている。この情報関連付け部１１は、図１に示す関連表検索部３の機能の中から分離された形で検索ログと図４に示した関連表作成機能とを有しており、詳細提示要求が入力されると、その検索ログから関連表を更新する。したがって、常に情報ニーズを反映した関連表が得られるようになっている。

【００４１】上記において、利用者の欲する情報が優先して出力される装置ならびに方法を説明したが、その方法はプログラムの形で記述して記録媒体に格納して再生可能に保管することができる。したがって、本発明は当該記録媒体をも技術範囲に含むものである。

【００４２】
【発明の効果】以上説明したように、本発明によれば、多くの利用者が使用した所定期間の検索ログを解析することにより得られた関連表に基づいて検索結果を自動的に順位づけして出力するため、利用者が欲する情報が優先して出力され、情報提供システムの操作性が極めて向上するという効果がある。

【図面の簡単な説明】

【図１】本発明を文書検索部分とともに構成した場合の一例を示すブロック図である。

【図２】本発明の処理を説明するフローチャートである。

【図３】検索語と詳細提示との関係を示す検索ログの一例を示す。

【図４】図３の検索ログから関連表を作成するブロック図である。

【図５】検索語と文書との対応関係を説明する図である。
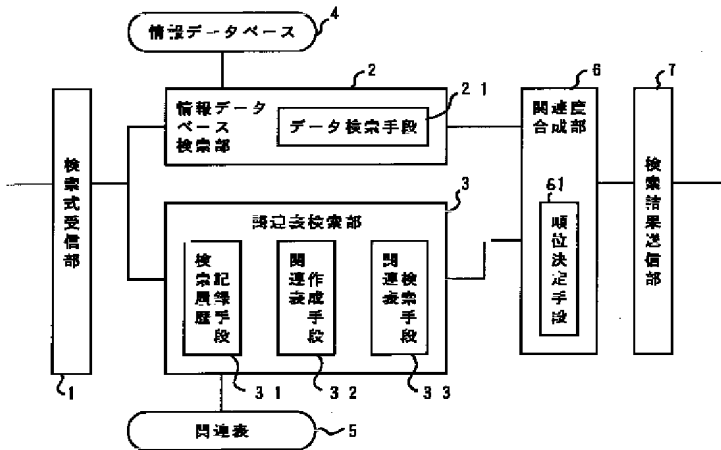
【図６】関連表の各行および各列に割り当てた記号を説明する図である。

【図7】本発明の別の実施形態の構成を示すブロック図である。

【符号の説明】
1　検索式受信部
2　情報データベース検索部
3　関連表検索部
4　情報データベース
5　関連表
6　関連度合成部
7　検索結果送信部
11　情報関連付け部

【図1】

【図3】

```
1998-11-17, 00:43:51, U12345, しし座*流星詳, 文書3
1998-11-17, 00:43:51, U12796, notaddress*登録, 文書128
1998-11-17, 00:43:56, U10195, しし座+獅子座, 文書3
1998-11-17, 00:44:00, U00285, ニセコ, 文書42
1998-11-17, 00:44:03, U62745, 宿泊施設+ホテル, 文書989
1998-11-17, 00:44:05, U12900, 論文*マーケティング, 文書1098
1998-11-17, 00:44:20, U12345, 獅子座流星群*撮影, 文書4096
...
```
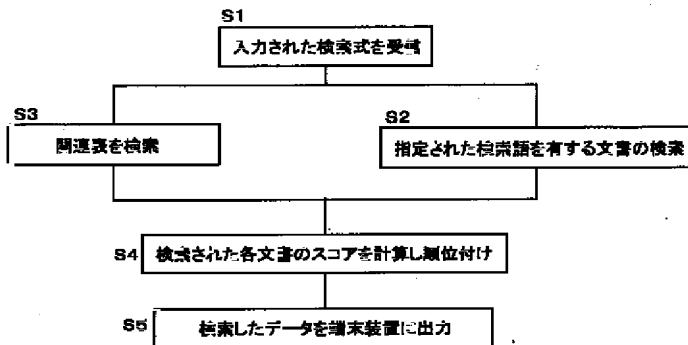
【図2】

【図4】

【図5】

文書 / 検索語	1	2	3	4	...	合計
:					...	
しし座			1		...	1
:					...	
流星群			1		...	1
:					...	
合計			2		...	

【図6】

文書 / 検索語	D_1	D_2	...	D_k	...	合計
X_1	$c_{1,1}$	$c_{1,2}$	...	$c_{1,k}$		$\sum_j c_{1,j}$
X_2	$c_{2,1}$	$c_{2,2}$	...	$c_{2,k}$		$\sum_j c_{2,j}$
:			...		...	
X_i	$c_{i,1}$	$c_{i,2}$	...	$c_{i,k}$		$\sum_j c_{i,j}$
:						
合計	$\sum_j c_{j,1}$	$\sum_j c_{j,2}$	...	$\sum_j c_{j,k}$	...	

【図7】